PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/17, 15/31, 15/62, C12P 7/64, C12N 1/19, 5/10, A61K 38/00, 38/28, 39/04, C07K 14/62, 14/35, C07H 13/06, G01N 33/74, 33/68

A1

(11) International Publication Number: **WO 95/22614**

(43) International Publication Date: 24 August 1995 (24.08.95)

(21) International Application Number: PCT/BR95/00010

(22) International Filing Date: 16 February 1995 (16.02.95)

(30) Priority Data:
PI 9400600-8 17 February 1994 (17.02.94) BR

(71) Applicants *(for all designated States except US)*: FINEP-FINANCIADORA DE ESTUDOS E PROJETOS [BR/BR]; Secretaria de Planejamento da Presidência da República, Praia do Flamengo, 200 - 13° andar, 22210-030-Rio de Janeiro, RJ (BR). ESCOLA PAULISTA DE MEDICINA [BR/BR]; Estabelecimento Federal de Ensino, Rua Botucatu, 740, São Paulo, SP (BR).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: CARDOSO DE ALMEIDA, Maria, Lucia [BR/BR]; Alameda dos Auêtis, 228, São Paulo, SP (BR). AMARAL DE CASTILHO VALAVICIUS, Beatriz [BR/BR]; Apartamento 31, Rua Urimonduba, 127, São Paulo, SP (BR). GOMES DE AMORIM FILHO, Antonio [BR/BR]; Rua Abílio Soares, 1003/81, São Paulo, SP (BR).

(74) Agent: DANNEMANN, SIEMSEN, BIGLER & IPANEMA MOREIRA; Rua Marquês de Olinda, 70, Botafogo, 22251-040-Rio de Janeiro, RJ (BR).

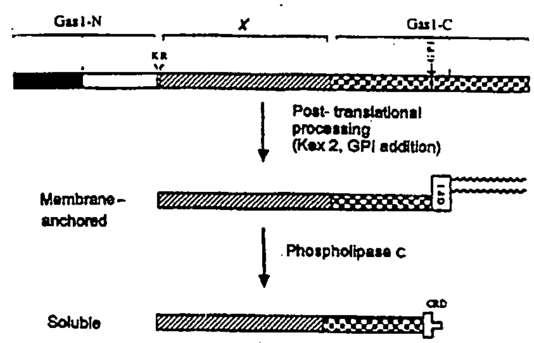
(81) Designated States: BR, CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PRODUCTION OF GLYCOSYLPHOSPHATIDYLINOSITOL-ANCHORED RECOMBINANT PROTEINS

(57) Abstract

Among others, a process is disclosed for producing a recombinant protein, or a precursor thereof, in cells of genetically modified eukaryotic microorganisms, especially *Saccharomyces cerevisiae.*, comprising the steps of biosynthesizing said protein or precursor by said cells and linking the endogenous glycosylphosphatidylinositol (GPI) to the C-terminal amino acid of the obtained protein or precursor, with the consequent anchorage of said protein or precursor to membranes of said microorganism by means of GPI; and selectively releasing the protein or precursor by methods which make use of intrinsic properties provided by the presence of GPI. GPI produced by the process can also be recovered. Recombinant proteins obtainable according to the invention include human insulin and *Mycobacterium leprae* 18kDa antigen.



Scheme of the precursor and of the GPI - anchored final product in S. cerevisiae, and of its release with phospholipase

"PRODUCTION OF GLYCOSYLPHOSPHATIDYLINOSITOL-ANCHORED RECOMBINANT PROTEINS"

The present invention relates to a process for expressing recombinant proteins, such as human insulin or **Mycobacterium leprae** 18kDa antigen, in genetically modified eukariotic microorganisms, especially in **Saccharomyces cerevisiae** yeast, in a form anchored to the plasma membrane by means of a glycosylphosphatidylinositol (GPI) anchor, and for selectively releasing them by methods making use of intrinsic properties provided by the presence of GPI, e.g. by treatment with GPI-specific phospholipase.

The purification of proteins of medical interest for use in diagnosis or treatment initially involves the large-scale obtention of the material from which the protein will be isolated. In many cases, this has become unfeasible for several reasons, such as costs and availability of the raw materials, these being in many cases human organs, which implies potential problems of contamination of the material with virus that cannot be totally eliminated during the purification process. For these reasons, in producing these proteins emphasis has been placed on the obtention of modified microorganisms that can express the products of interest. For instance, bacteria such as **E.coli** can be modified in such a manner that they become capable of synthesizing human proteins in large amounts. Yeasts such as **S.cerevisiae** and cells of mammals maintained in culture can also be altered in order to

produce the peptides of interest. The advantages in these cases are the low cost of the raw material, the reproducibility in the obtention of large amounts of the product and the harmlessness of the systems employed. Another great advantage of these systems is the possibility of altering the sequence of the proteins being expressed, in order to adapt them to the needs of the product; these alterations, involving residues of certain amino acids, in theory can result in the obtention of products that are more specific, have a higher activity or exhibit additional functions not presented by the original product. For instance, the usually employed insulin is obtained from the pancreas of oxen or pigs. However, these animal insulins differ from human insulin by 3 and 1 amino acid residues, respectively, and this minor variation accounts for undesired immunological reactions. Therefore, the large-scale production of recombinant human insulin represents a very important step in the therapy of diabetes. Another important advantage of the obtention of polypeptides by recombinant DNA technology is the possibility of modifying the product in order to facilitate its purification. For instance, additional residues that can be recognized and linked to a specific substance such as an antibody, lecitins or protein A, can be introduced in the recombinant protein, thus facilitating the large-scale purification of the product by affinity.

The yeast **Saccharomyces cerevisiae** has become an important tool in the production of recombinant proteins and is often preferred to **E.coli** since the latter has toxic substances associated with its wall, which implies extensive purification processes and tests of the products synthesized in this bacterium. The yeast, on the other hand, has been used for centuries in the manufacture of foodstuff and is harmless to human beings.

The recombinant proteins can be produced by these hosts and remain inside the cells. Alternatively, these proteins can be secreted into the culture medium. Each method has its advantage and, depending on the type of protein to be expressed, one or the other is chosen. In general, for proteins which are usually secreted by the cells that express them naturally, the secretion way is chosen for their production on a

large scale. Naturally intracellular proteins are usually produced in these organisms also in the intracellular form. However, although mimicking the native location, in many cases and for reasons which cannot be determined beforehand, one cannot succeed in producing them on a large-scale and it is necessary to test several forms of biosynthesis. Many proteins, when expressed intracellularly in large amounts, form inclusion bodies constituted by proteins aggregated in an inactive manner. Although the purification of these proteins is facilitated by their being in an insoluble form, their use is limited because this requires steps of denaturation and renaturation **in vitro**, which do not always allow the obtention of proteins with their natural and active conformation and, besides, result in low yields of the correctly folded product.

The process of purifying the recombinant product also defines the type of expression of the protein. Thus, for instance, secreted proteins will be in a diluted form in the culture medium, which should be processed after elimination of the cells by centrifugation or filtration. This implies a concentration and purification of the product starting from hundreds or thousands of litres, requiring quite sophisticated and complex procedures. On the other hand, few proteins are secreted normally, for which reason the product of interest will comprise a large portion of the total protein of a supernatant of these cultures. Intracellularly produced proteins are concentrated by merely collecting the cells. In this case, the protein of interest will constitute a small amount of the total protein and, therefore, a number of purification steps are involved.

The present invention provides an alternative way of expressing recombinant proteins in **S.cerevisiae**, by which the recombinant product will be associated with the plasma membrane by means of a GPI anchor. The use of this method for expressing human insulin and the **M.leprae** 18kD antigen is exemplified hereinbelow. However, this technology is not limited to these examples. Other polypeptides can also be expressed in this form. The advantages in using this technology will be apparent from the following description.

All the eukaryotic organisms have proteins that are

normally associated with the plasma membrane by means of glycosylphosphatidylinositol (GPI) covalently bound to the C-terminal amino acid of the polypeptide chain. GPI, which is a structure quite well conserved throughout the evolutionary scale, contains ethanolamine, phosphate, mannose, occasionally galactose and necessarily a non-acetylated glucosamine, which is in glycosidic linkage with the inositol ring of a phosphatidylinositol molecule (Ferguson & Williams, 1989, Ann. Rev. Biochem. 59:285; Thomas et al., 1990, Biochemistry **29** :5413). This structure is co-translationally added at the level of the endoplasmic reticulum in a process of which the rapidity suggests that the nascent protein receives the already preformed GPI anchor by means of enzymatic steps not yet totally characterized (Doering et al., 1989, J. Biol. Chem. 265:611). The precursors of the GPI-anchored proteins contain a hydrophobic sequence of from 15 to 30 amino acids at the C-terminal, which is eliminated during the processing, generating a new C-terminus, to which the GPI structure is added. Thus, both transmembranic and secreted proteins can be converted into their respective GPI-anchored variants by adding this signal sequence to their respective biosynthetic precursors (Caras et al, 1987, Science 238:1280; Crise et al., 1989, J. Virol. 63 :5328; Tykocinsky et al., 1988, Proc. Natl. Acad. Sci. USA **85** :3555). Although there is no absolute consensus regarding the primary structure of the terminal peptide, the analysis of the sequences that flank the site of addition of GPI indicates that three aspects seem important in the process, in both mammals and protozoa: a hydrophobic sequence of a minimum size, the absence of possible cytoplasmatic domains and a somewhat indefinite pattern of amino acid sequences or conformation around the site of cleavage (Low, 1990, Biochim, Biophys. Acta **988** 427; Ferguson & Williams, 1989, op. cit.; Caras & Weddell, 1989, Science **243** :1196).

The GPI-anchored proteins are usually released from the membrane when treated with inositol-specific phospholipase, which treatment, upon hydrolysing phosphatidylinositol, breaks the domain of interaction with the membrane (Low, 1990, op.cit.). Several types of

phospholipase C (PLC) and D capable of solubilizing GPI-containing molecules in vitro have been described. Some of these types, such as the phospholipases C of bacterial origin, have a broad specificity and hydrolyse phosphatidylinositol and GPI, while others, from mammals or trypanosomes, seem to be highly specific for GPI (Low, 1990, op. cit.). In their native condition GPI-containing glycoproteins exhibit an amphiphilic character and, after solubilization by PLC, acquire a hydrophilic character (Ferguson & Williams, op. cit.). Based on this difference of behaviour, easily induced by enzymatic treatment, several purification techniques based on phase separation in Triton X-114 (Bordier, 1981, J. Biol. Chem. **256** :1604; Cardoso de Almeida & Turner, 1983, Nature **302** :349, and charge displacement electrophoresis, Toutant et al., 1989, Eur. J. Biochem. 180:503) have been employed for obtaining and characterizing glycoproteins with this post-translational modification.

Besides the in vitro enzymatic treatment, the release of GPI-anchored proteins can also be mediated **in vivo** by means of the co-expression of phospholipases (Scallon et al., 1992, Bio/Technology **10**:**500**). The genes coding for some phospholipases, including GPI-specific phospholipase C, have already been cloned and sequenced; the expression of these activities in cells of interest is, therefore, possible.

Concomitantly with the solubilization by PLC the released protein exposes inositol 1,2-cyclic phosphate in the structure of its hydrolysed GPI, which is a structural marker known as "cross-reacting determinant" (CRD), and an epitope classically recognized as an antibody (anti-CRD). This polyclonal antibody has the property of discriminating the entire structure of GPI from the structure hydrolysed by phospholipases C and further allows the selective purification of GPI-containing molecules by immunoaffinity. (Cardoso de Almeida & Turner, 1983, op. cit.; Zamze et al., 1988, Eur. J. Biochem. 176 :527).

Consequently, the GPI moiety of an anchored protein can also be obtained by chemical or enzymatic treatments, and free GPI's can be purified by selective extraction with organic solvents (Orlean et al, 1994, Brazilian J. Med. Biol.

Res. Vol. 27, page 145).

In **S. cerevisiae** only one protein is predominantly solubilized with phospholipase C and visualized with anti-CRD antibody (Conzelmann et al., 1988, EMBO J. 7 :2233). The function of this 125 kDa protein, called Gas1, is still unknown. The molecular characterization of the gene which codes for it (**GAS1**) indicates that this protein is not necessary for the normal growth of the yeast.

In view of the foregoing, any protein can receive a GPI-anchor and consequently remain linked to the plasma membrane, if the appropriate signal sequence is added to its carboxyl terminus. As it will be exemplified hereinafter, in **S.cerevisiae** lacking the Gas1 endogenous protein, the recombinant product containing the GPI-anchor will practically be the only protein to be recognized and cleaved by phospholipase. Therefore, this method represents a unique advantage of releasing the product in a highly specific and selective manner. In addition, the recombinant product will contain an epitope (CRD) which can be used in its final purification by immunoaffinity.

Human insulin, commercially available for the treatment of insulin-dependent diabetes is derived, at present, from three alternative sources, namely: a) chemically modified animal insulin (Markussen, J., 1980, US Patent 3,433,898); b) recombinant human insulin, produced in **Saccharomyces cerevisiae** in secreted form (Thim et al., 1987, FEBS Letters **212** :307), and c) recombinant human insulin produced in **E. coli** in the form of intracellular pro-insulin, which after purification is enzymatically cleaved with tripsin and carboxypeptidase B, providing human insulin (Williams et al., 1982, Science 215:687; Sung et al., 1986, Proc. Natl. Acad. Sci., 83:561). The present invention is exemplified in this alternative form of expressing human insulin in the yeast **S.cerevisiae**, in which the insulin molecule will be originally anchored via GPI to the cellular membrane, which allows the facilitated purification of the product by means of the treatment of the cells with GPI-specific phospholipase C.

The **Mycobacterium leprae** 18kDa protein is one of the main targets of the cellular immune response in individuals

infected by this bacillus (Booth et al., 1988, J. Immunol. 140:597). **M.leprae** is at present a problem of underdeveloped and developing countries, causing leprosy in approximately 15 million people (WHO Bulletin, 1988, World Health Organization/Switzerland). Due to the difficulty in cultivating it, little is known about the biology, biochemistry or immunology of this disease. It was suggested that the 18kDa protein might represent a potential vaccine against this disease (Young et al., 1988, Proc. Natl. Acad. Sci. USA 85:4267). With a view to making a indepth study of the cellular and humoral immune response to this antigen, many attempts have been made to obtain its expression and purification on a large scale in **E. coli**, but without success due to the formation of insoluble aggregates, difficulty in the purification and proteolytic cleavage of the recombinant protein. The problem was appearently solved by its expression in **S. cerevisiae**, in both intracellular and extracellular form (Booth et al., 1988, Immunol. Lett. 19 :65; Piestun, 1992, Master Dissertation, Dept. Immunol., ICB, YSO). The present invention represents an alternative for obtaining a **M. leprae** 18kDa protein in membrane-anchored form in **S. cerevisiae**. This membrane-anchored form can offer advantages over the soluble forms of this protein, or of other antigens, for vaccination using, for instance, liposomes covered with this GPI-anchored protein. In addition, this anchored protein can be used in agglutination assays, which are admittedly the most economical serological test for infectious diseases, especially in less developed countries. Besides, a GPI-containing immunogen can elicit an immune response modulated by parameters that, although not yet well-known, can favour the protection (Schofield & Harckett, 1993, J. Exp. Medicine **177** :145).

Thus, the present invention represents an alternative in the way of expressing recombinant proteins in **Saccharomyces cerevisiae**, with many practical possibilities, comprising:

a) the protein will contain a carboxy-terminal moiety capable of locating it in the cytoplasma membrane;

b) said localization will be due to the presence of a structure called glycosylphosphatidylinositol (GPI) anchor,

the addition of which to the protein is mediated by the carboxy-terminal sequence referred to under a);

c) the carboxy-terminal sequence will be derived from the **S.cerevisiae** Gasl protein;

d) the expressed protein can be any protein component usually found in eurokaryotic or prokaryotic organisms, such as hormones, surface antigens, secreted peptides, cytoplasmatic proteins etc;

e) the expressed protein can be human insulin, or the **Mycobacterium leprae** 18kDa antigen;

f) the **S.cerevisiae** strain host can be one lacking the Gasl endogenous protein;

g) the GPI-anchored protein can be detected by the presence of the epitope CRD after treatment with GPI-specific phospholipase C, or by the alteration of its amphiphilicity after treatment with GPI-specific phospholipase;

h) the GPI-anchored protein can be released specifically from the insoluble fraction of membranes, from a extract of these yeasts, prepared by breaking with glass beads, by the action of GPI-specific phospholipase.

The present invention also comprises a method of obtaining yeast strains expressing human insulin or the **Mycobacterium leprae** 18kDa antigen in membrane-anchored form, which method comprises:

a1) chemical synthesis of the insulin gene, coding only for chains B and A, these chains being linked by dibasic residues; or

a2) enzymatic synthesis of the sequence coding for the **Mycobacterium leprae** 18kDa antigen;

b) construction of vectors suitable for expressing GPI-anchored proteins in **S.cerevisiae**;

c) insertion of the sequences coding for insulin or for the 18kDa antigen in these vectors; and

d) introduction of the plasmids obtained in c) in yeast strains lacking the endogenous protein Gasl.

This invention also comprises detecting the production of insulin or of the 18kDa antigen in the cells transformed with the plasmids, by means of immuno-blots from SDS-PAGE or directly from colonies.

This invention further comprises detecting the presence of GPI anchor in the proteins insulin and 18kD, by means of immuno-blots with antibody anti-CRD after treatment of the filters with PLC.

Gas1 being pratically the only GPI-anchored protein present in the yeast cell which is sensitive to hydrolysis by GPI-specific phospholipase C, the present invention provides a highly selective method of purifying proteins.

Thus, an object of the present invention is to provide a process for producing a recombinant protein, or a precursor thereof, in cells of genetically modified eukariotic microorganisms, characterized by comprising the steps of:

a) biosynthesizing said protein or a precursor thereof by the cell of said microorganism and linking the endogenous glycosylphosphatidylinositol (GPI) to the C-terminal amino acid of the obtained protein or its precursor, with the consequent anchoring of said protein or precursor to membranes of said microorganism by means of GPI; and

b) selectively releasing the protein or its precursor obtained in step a) by methods making use of intrinsic properties provided by the presence of GPI.

A specific embodiment of this process is characterized in that said eukaryotic microorganism is a genetically modified **Saccharomyces cerevisiae.**

Other eukaryotic microorganisms that are also contemplated herein are **Dictyostelium discoideum,** trypanosomes and other yeasts such as **Pichia pastoris** and **Hansenula polymorpha.**

Another specific embodiment of this process is characterized in that said eukaryotic microorganism is **Saccharomyces cerevisiae** genetically modified so as not to produce endogenous GPI-anchored proteins.

Another specific embodiment of this process is characterized in that said eukaryotic microorganism is **Saccharomyces cerevisiae** genetically modified so as not to produce the **Gas1** endogenous protein.

Another specific embodiment of this process is characterized in that said recombinant protein is human insulin.

Another specific embodiment of this process is char-

acterized in that said recombinant protein is **Mycobacterium leprae** 18kDa antigen.

Another specific embodiment of this process is characterized in that the selective release of the protein or its precursor obtained in step a) is carried out by enzymatic treatment.

Another specific embodiment of this process is characterized in that the enzymatic treatment is carried out with PI- or GPI-specific phospholipase.

Another specific embodiment of this process is characterized in that the enzymatic treatment is carried out with PI- or GPI-specific phospholipase of C or D specificity.

In a more specific embodiment of this process, one works with PI- or GPI-specific phospholipase C, obtaining hydrolysed protein or its precursor, which, in its structure, presents inositol 1,2-cyclic phosphate (epitope CRD), which is susceptible of being selectively purified by immunoaffinity with anti-CRD antibodies.

Another specific embodiment of this process is characterized in that the enzymatic treatment is carried out with proteases.

Another specific embodiment of this process is characterized in that the selective release of the protein or its precursor obtained in step a) is carried out by chemical treatment.

Another specific embodiment of this process is characterized in that the chemical treatment is carried out by nitrous deamination.

Another specific embodiment is characterized in that the chemical treatment is carried out with a base.

Another specific embodiment of this process is characterized in that the selective release of the protein or its precursor obtained in step a) is carried out by treatment with neutral detergents.

In another specific embodiment of this process, this treatment is carried out with neutral detergents of the n-octylglucopyranoside type.

Another specific embodiment of this process is characterized in that the selective release of the protein or its

precursor obtained in step a) is carried out by co-expression of a PI- or GPI-specific phospholipase in the cell of said microorganism.

It is also an object of this invention to provide a process for the production of glycosylphosphatidylinositol (GPI) and a recombinant protein in cells of genetically modified eukaryotic microorganisms, characterized by comprising the steps of:

a) biosynthesizing said protein or a precursor thereof and biosynthesizing GPI or a precursor thereof by the cell of said microorganism, and linking the endogenous glycosylphosphatildylinositol (GPI) to the C-terminal amino acid of the obtained protein or its precursor, with the consequent anchoring of said protein or its precursor to membranes of said microorganism by means of GPI;

b) selectively releasing the obtained protein or its precursor by methods making use of intrinsic properties provided by the presence of GPI; and

c) selectively releasing GPI.

In a specific embodiment of this process, said eukaryotic microorganism is a genetically modified **Saccharomyces cerevisiae.**

Other eukaryotic microorganisms that are also contemplated here are **Dictyostelium discoideum**, trypanosomes and other yeasts, such as **Pichia pastoris** and **Hansenula polymorpha**.

In a specific embodiment of this process, said eukariotic microorganism is a **Saccharomyces cerevisiae** genetically modified so as not to produce GPI-anchored endogenous proteins.

In a specific embodiment of this process, said eukaryotic microorganism is a **Saccharomyces cerevisiae** genetically modified so as not to produce the endogenous **Gasl** protein.

In a specific embodiment of this process, said recombinant protein is human insulin.

In another specific embodiment of this process, said recombinant protein is **Mycobacterium leprae** 18kDa antigen.

As a specific embodiment of such process the selec-

tive release of a protein or its precursor or GPI or its precursor obtained in step a) is obtained by enzymatic treatment. This treatment can make use of phospholipases with PI- or GPI-specificity, and specially with PI- or GPI-specific phospholipases of C or D specificity. After rupture of yeast the resulting particulate fraction containing the membranes is resuspended in a minimum volume of 0.1% Triton X-114 in 50 mM Tris.HCl pH 7.4 with 150 mM NaCl. Then, 1600 U/ml of the phospholipase C either from **Trypanosoma brucei** (GPI-specific) or from **Bacillus thuringiensis** (PI-specific) is added and the mixture, after gentle mixing, is incubated for 1 hour at 30°C. Alternatively, the yeast membranes containing the GPI-anchored recombinant protein can be resuspended in 50 mM Tris.acetate pH 5.4 and 4000 U/ml of GPI-specific phospholipase D from human or rat is added and the reaction mixture incubated for 1 hour at 30°C. At the end of these incubations either the supernatant after centrifugation or the whole incubation mixture will contain the target protein or the precursor in its hydrolysed form which presents in its structure the inositol 1,2-cyclic phosphate (anti-CRD epitope), which can be selectively purified by immunoaffinity with anti-CRD antibodies. Antibody with this specificity can be obtained from rabbits immunized with ILTat 1.21 mf VSG from **T.brucei** after its conversion to its soluble form ILTat 1.21 sVSG by the action of 1600 U/ml of **T. brucei** phospholipase C in 10 mM Tris.HCl pH 7.4 containing 0.05% Triton X-114 and 0.05% n-octylglucopyranoside for 3 hours at 30°C. New Zealand rabbits are immunized with 100 micrograms of ILTaT 1.21 sVSG produced as described above, in form of an emulsion with 0.5 ml of 137mM NaCl, 2.7 mM KCl, 4.5mM Na_2HPO_4, 1.5 mM KH_2PO_4pH 7.4 (PBS) and 0.5 ml of Freund's complete adjuvant commercially available. Subsequent boosts at monthly intervals are made with 25 micrograms of ILTaT 1.21 sVSG in 0.5 ml of PBS in Freund's incomplete adjuvant. Rabbits are bled exactly 2 weeks after each boost. The serum from several rabbits is pooled and applied to a column in which MITat 1.5 VSG has been immobilized on Sepharose CL 4-B (Pharmacia) using standard protocols. The antibody population bound to the column is eluted with 0.1M glycine pH 2.4 and each ml of fraction is

collected on ice and immediately neutralized with 90 microlitres of 1M Tris.HCl pH 7.4. The elution of the antibody is followed by following optical density at 280 nm, and all the fractions with ODs above 0.1 are pooled and applied to a column where MITat 1.6 mf VSG has been immobilized on Sepharose-CL-4B (Pharmacia) using standard protocols. The fraction of antibody which does not bind to the column is the so called anti-CRD antibody which will react with GPI-anchored proteins hydrolysed by phospholipase C but not by phospholipase D. This antibody is then immobilized on Sepharose CL-4B and this resin can be used several times to purify, by affinity, phospholipase C-solubilized GPI components.

The enzymatic treatment for release of GPI-anchored proteins can also be performed by specific proteases provided that an unique site for its action is carefully planned at the design of the recombinant protein. In the case of the constructs mentioned in this patent application a trypsin site was inserted exactly at the fusion sites. Accessibility of the protease to the site might depend on folding of each recombinant protein. In this case the released products will not contain the CRD epitope since the GPI is left inserted in the plasma membrane.

Still in a further embodiment of this process, the selective release of the protein or its precursor or GPI or its precursor obtained in step C is obtained by chemical treatment with 100 mM NaOH for 1 hour at 37°C in the case of di-acylglycerol-based anchors or alternatively by addition of 48% hydrofluoric acid. Solvent extraction can be performed by treating each ml of membranes with 6.9 ml of a mixture of chloroform and methanol (1:1), this being an extraction in chloroform : methanol : water (10:10:3 v/v/v). After extraction of GPI or precursors for 2 hours at room temperature, the insoluble material is centrifuged at 4,000 g for 10 minutes and the supernatant dried under a N_2 stream. This material is then re-suspended in 100 microlitres of water-saturated 1-butanol and the salts removed by addition of 50 microlitres of water. The aqueous phase is again re-extracted with 100 microlitres of

water-saturated 1-butanol. The same procedure is still repeated for a third time and the pooled organic phases washed twice with 50 microlitres of water-saturated 1-butanol and then dried under N_2.

A further specific embodiment of this process of releasing protein or its precursor obtained in the step c) requires treatment with neutral detergents such as n-octylglucopyranoside. In this type of treatment the membranes containing the product are extracted with 0.5% of the detergent in PBS and centrifuged at 100,000 g for 1 hour at 4°C.

Still another specific embodiment of this process of releasing the protein or its precursor obtained in step c) consists of co-expressing a PI- or GPI- specific phospholipase in the same microorganism.

It is also an object of this invention to provide a process for the production of glycosylphosphatidylinositol (GPI) in cells of genetically modified eukaryotic microorganisms, characterized by comprising the steps of:

a) biosynthesizing a protein or a precursor thereof and biosynthesizing GPI or a precursor thereof by the cell of said microorganism and possibly linking the endogenous glycosylphosphotidylinositol (GPI) to the C-terminal amino acid of said protein or its precursor obtained in step a), with the consequent anchoring of said protein or its precursor to membranes of said microorganism by means of GPI; and

b) selectively releasing GPI.

In a specific embodiment of this process, said eukaryotic microorganism is genetically modified **Saccharomyces cerevisiae.** Other eukaryotic microorganisms that are also contemplated here are **Dictyostelium discoideum**, trypanosomes and other yeasts, such as **Pichia pastoris** and **Hansenula polymorpha.**

In a specific embodiment of this process, said eukaryotic microorganism is **Saccharomyces cerevisiae** genetically modified so as not to produce GPI-anchored endogenous proteins.

In a more specific embodiment of this process, said

eukaryotic microorganism is a **Saccharomyces cerevisiae** genetically modified so as not to produce the **Gasl** endogenous protein.

It is also an object of this invention to provide a process for the obtention of **Sacchromyces cerevisiae** yeast cells capable of expressing a recombinant protein, characterized by comprising the steps of:

a) providing the gene which codes for said recombinant protein or its precursor so as to be anchored via GPI;

b) inserting the gene obtained in step a) in an adequate vector;

c) introducing the vector containing the gene which codes for said recombinant protein or its precursor in a **Saccharomyces cerevisiae** cell; and

d) propagating the cell obtained in step b).

In a more specific embodiment of this process, said recombinant protein is human insulin.

In another more specific embodiment of this process, said recombinant protein is **Mycobacterium leprae** 18kDa.

In another more specific embodiment of this process, the **Saccharomyces cerevisiae** cell used lacks the capacity of producing the **Gasl** endogenous protein.

It is also an object of this invention a yeast cell characterized by being genetically modified so as to express a GPI-anchored recombinant protein.

In a more specific embodiment, this yeast cell is **Saccharomyces cerevisiae.** In still a more specific embodiment, this yeast cell is characterized by not expressing the **Gasl** endogenous protein.

In a further more specific embodiment, this yeast cell is characterized by the capacity of raising the levels of GPI biosynthesis or of producing GPI with determined properties.

In a further more specific embodiment, this yeast cell can be characterized by the fact that said recombinant protein is human insulin or **Mycobacterium leprae** 18kDa antigen.

It is also an object of this invention a nucleotide

sequence characterized by coding for a recombinant protein or its precursor susceptible of being anchored by GPI.

Also an object of this invention is to provide a nucleotide sequence characterized by coding for:

(i) the C-terminal moiety of the **Gas1** protein of **Saccharomyces cerevisae;**

(ii) a protein or its precursor; and

(iii) the N-terminal moiety of the **Gas1** protein of **Saccharomyces cerevisiae.**

In specific embodiments, this nucleotide sequence can be characterized in that said protein is human insulin or in that such a protein is **Mycobacterium leprae** 18kDa antigen.

It is also an object of this invention a culture medium characterized by containing cells in accordance with this invention described in the preceding paragraphs.

It is a further object of this invention to provide a medicament or vaccine characterized by containing a recombinant protein obtained by the processes of this invention described above, or by containing a recombinant protein whose precursor has been obtained by one of the processes of this invention described above. Compositions containing recombinant human insulin or **Mycobacterium leprae** 18kDa antigen are specifically contemplated here.

In order to help in the interpretation of the meaning of the present invention, some terms used in the text are described below.

Coding sequence - A sequence of DNA which, when transcribed and translated, results in the formation of a polypeptide.

Gene - A region of the genome comprising the coding sequence and sequences responsible for the control of its expression, that is to say, transcription and translation.

Secretion signal sequence - A sequence of hydrophobic amino acids present in the amino-terminal moiety of a polypeptide, which has the function of directing this polypeptide to the endoplasmic reticulum and, consequently, in the absence of any other signal, to the outer environment.

A Sequence Signalizing the Addition of a GPI Anchor - A sequence of hydrophobic amino acids present in the

carboxy-terminal moiety of a polypeptide signalizing the proteolytic cleavage in a given residue, to which the GPI anchor will be added.

GPI Anchor - A glycolipid linked to the C-terminal amino acid of proteins intended for being anchored to the cellular lipidic membrane.

In order to construct fusions of a given polypeptide and a GPI-anchor domain, DNA coding for the C-terminal 30-50 residues of a protein that usually contains this anchor is ligated to DNA coding for the polypeptide in question. This fusion is made in the C-terminal moiety of said polypeptide. The C-terminal 10 - 20 hydrophobic residues will be processed after translation and eliminated from the mature protein. The construction of this fusion is accomplished by those skilled in the art by means of routine techniques. For instance, DNA coding for the region signalizing the addition of an anchor can be synthesized **in vitro** or isolated from a genomic DNA or cDNA.

The host for the expression of these anchored proteins is **S. cerevisiae** especially modified, non-coding for Gas1. Gas1 is the main or the only GPI-anchored protein which is susceptible to the action of GPI-specific phospholipase C. Thus, when strains are used from which the gene **GAS1** has been deleted, the treatment of cellular extracts of this strain expressing the specific GPI-anchored protein will release only the recombinant product essentially free from other contaminants.

In order to obtain these products in the mentioned hosts, the DNA coding for the product is introduced in these hosts by several techniques, for instance, transformation, electroporation, transfection etc. However, the mere introduction of the DNA coding for the product in these cells is not enough for them to start synthesizing it. It is necessary for this DNA to be capable of being transcribed into a messenger RNA molecule; for this purpose the DNA has to be coupled to a promoter sequence, which is a deoxynucleotide sequence specifically recognized by the transcription machinery of the cell in question. Once this sequence is recognized, the RNA polymerase starts to synthesize the messenger RNA using the

DNA molecule in question as a template. The promoter sequences are specific for each organism. Thus, if the protein in question is produced in **S. cerevisiae** the coder DNA should be preceded by a promoter sequence which is recognized by the **S. cerevisiae** RNA polymerase. Once the messenger RNA is synthesized, it can be translated into a protein molecule. It is also important for the translation signals to be present in the sequence introduced in the cell.

The sequence coding for the protein that is to be obtained can be obtained by several cloning methods known to those versed in the art. A DNA segment can be obtained directly from the chromosomal DNA, can be synthesized from the messenger RNA or can be chemically synthesized.

This DNA should then be inserted downstream of a promoter sequence. Several promoters can be used for expression in **S. cerevisiae** and are in the public domain. For instance, the promoter pGAL1, derived from the gene **GAL1** or the promoter **ADH2**, derived from the gene coding for alcoholic dehydrogenase gene can be used.

The type of promoter is chosen depending on the system to be employed or on the process or on the product. There are constitutive promoters and, therefore, in these cases, the cell will produce the protein in question continuously, which can be favourable in cases of co-expression of GPI-specific phospholipase. In other cases, it is preferred that the cell produces the protein in question only when it is programmed for this purpose, and in such cases the so-called inducible promoters are used, that is to say, promoters that are specifically activated when their functioning is required, usually at the end of the fermentation process.

In order for the organism to produce the protein in question, it is not sufficient for its coding sequence to be coupled to a promoter and this assembly to be introduced in the organism. This assembly has to be maintained stable within that organism, avoiding its spontaneous loss. For this purpose, this assembly is incorporated in plasmids, which are generally circular DNA molecules, maintained independently inside the cells. Several types of plasmids are known and in the public domain. General characteristics of plasmids used as

vectors are: a) being capable of autonomously replicating within the host; b) having a selectable marker, that is to say, a gene that is essential for the survival of the host. Typical selectable genes are: a) for selection in bacteria, they should preferably impart resistance to antibiotics, for instance ampicillin or tetracycline; b) for selection in **S. cerevisiae**, they should complement auxotrophic defficiencies of the host, that is to say, they should contain genes which allow the synthesis of essential components absent from the culture medium.

In order to simplify the description of the examples, certain methods will be mentioned by their routine nomenclature.

"Plasmids" are designated by a "p" followed by capital letters and/or numbers. The starting plasmids described herein are commercially available, are publicly available without restrictions, or can be constructed from such available plasmids in accordance with published procedures. In addition, other equivalent plasmids are known to those skilled in the art.

"Digestion" of DNA refers to the catalytic cleavage of the DNA with an enzyme that only acts at certain locations of the DNA. These enzymes are called restriction enzymes, and the specificity sites are called restriction sites. The enzymes used herein are commercially available, and their reaction conditions, co-factors and other requirements are those supplied by the manufacturers. Restriction enzymes are designated by three letters followed by a number. In general, about 1 microgram of plasmidial DNA or DNA fragment is used with about 2 units of the enzyme, in about 20 microliters of buffer solution, as specified by the manufacturer. Incubation proceeds for about 1 hour at 37°C, or as specified by the manufacturer. After incubation, protein is extracted with phenol and chloroform, and the digested DNA is recovered from the aqueous phase by precipitation with ethanol. Infrequently the plasmidial DNA can be dephosphorylated by treatment with the bacterial alkaline phosphatase enzyme, which removes the terminal 5' phosphates, preventing the circularization of the plasmid during the reaction with the enzyme ligase T4, which

would prevent the insertion of another DNA fragment in that restriction site.

"Filling" or "blunting" refers to the procedure by which the cohesive termini left by a restriction enzyme are filled with deoxynucleotides by the action of the DNA polymerase I enzyme (Klenow fragment), becoming blunt. The filling of the termini allows the linking of any fragment thus treated with another blunt-ended DNA. Typically, the reaction consists of the incubation of the target DNA with a buffer according to the specification of the manufacturer, in the presence of 8 units of DNA polymerase I (Klenow) and 250 microM of each of the four deoxinucleotides. The incubation proceeds for 30 minutes at 37°C, and is followed by extraction with phenol and chloroform, and precipitation with ethanol.

"Isolation of DNA fragment" refers to the separation of the products of a digestion with a restriction enzyme by electrophoresis in agarose gel, identification of the fragment of interest by comparing its migration with standards of known molecular weights, removal of the gel portion containing the fragment of interest and separation of the DNA from the gel, according to well-known techniques. For instance, see Sambrook et al., 1989).

"Ligation" refers to the process of forming a phosphodiester bonds between two double-stranded DNA fragments (Sambrook). The ligation reaction consists of the incubation of 0.5 microgram of the two DNA fragments in approximately equimolar amounts, in the presence of 10 units of the DNA T4 ligase enzyme, in the presence of a buffer as specified by the manufacturers.

"Transformation" refers to the method of introducing DNA in an organism, so that this DNA replicates extrachromosomally or integrated in the chromosome. The method of transforming E. coli is the one described by Mandel et al., 1970; J.Mol.Biol. 53,154. The method of transforming S. cerevisiae is the one described by Ito et al., 1983, J. Bacteriol. 153; 163.

"Preparation of DNA of the transformants" refers to the method of isolating plasmidial DNA from microbial cultures. The method employed for bacteria can be that of

alkaline lysis in the presence of SDS, as described in Sambrook et al., 1989, **Molecular Cloning**, Laboratory Manual.

"Oligonucleotides" are short polymers of deoxynucleotides, which are chemically synthesized by well-known methods.

"Western blots" or "immuno-blots" refer to the method of identifying proteins separated in SDS-acrylamide gel by the use of specific antibodies. The cellular extracts or fractions thereof are applied to polyacrylamide denaturating gel (Laemmli, 1970, Nature 227:680) and subjected to electrophoresis. The separated proteins are transferred to nitrocellulose filters and subsequently subjected to reaction with antiserum or specific monoclonal antibodies.

"PCR Reaction" refers to the enzymatic amplification of a double-stranded DNA chain starting from two oligonucleotides complementary to the two 5' ends of the template molecule. The DNA to be amplified can be genomic DNA or DNA present in a plasmid. The reaction uses the enzyme Taq polymerase and is carried out as specified by the manufacturer, in a thermal cycle apparatus, for approximately 40 cycles.

"Sequencing" of DNA refers to the determination of nucleotides present in a DNA chain and can be carried out by the enzymatic method with the enzyme Sequenase, as specified by the manufacturer, and the DNA chains generated are separated in polyacrylamide-urea denaturating gel and visualized by autoradiography, as described in Sambrook et al., op.cit..

Nucleotides are designated by the standard letters, A being adenine, T thymine, C cytosine and G guanine.

Aminoacids are represented by the designations of a letter, wherein A is alanine, R is arginine, N is asparagine, D is aspartate, C is cysteine, Q is glutamine, E is glutamate, G is glycine, H is histidine, I is isoleucine, L is leucine, K is lysine, M is methionine, F is phenylalanine, P is proline, S is serine, T is treonine, W is tryptophan, Y is tyrosine, V is valine.

"GPI extraction" means the recovery of this glycoinositolphospholipid by using a series of organic solvent/aqueous extractions which selectively separate it from

proteins, DNA, lipids and phospholipids. The products of this extraction can be analysed using standard processes of thin layer chromatography or reversed phase thin layer chromatography. Occasionally it is also appropriate to fractionate the products on an Octyl-Sepharose column (Pharmacia), using appropriate gradients of 5-80% of 1-propanol.

The scheme and the figures will now be described in detail, to which reference will be made in the examples of use of this invention.

Figure 1 - Scheme of the Precursor and of the GPI-anchored Final Product in S. cerevisiae, and of its Release by Phospholipase. Any protein "X" can be coupled to the N-terminal region of Gas1 and to the C-terminal region of Gas1. The post-translational processing during transport to the endoplasmic reticulum will allow the elimination of the N-terminal sequences of Gas1, through signal peptidase and the endopeptidase Kex2, which recognizes and cleaves in dibasic residues, and the coupling of the GPI anchor. This polypeptide will remain anchored to the membrane. Through the action of phospholipase C, the anchor will be hydrolysed, and the polypeptide can be released from the membrane, with the simultaneous creation of the CRD epitope.

Figures 2A and 2B. Strategy of Construction of Vectors for Expression of GPI-anchored Insulin (2A) and GPI-anchored 18kDa Protein (2B) in S. cerevisiae. The plasmids of the series pBY constructed here are derived from YEp352, and are bifunctional, that is to say, they replicate and can be selected in both **E.coli** and **S.cerevisiae.** The other plasmids are replicable only in **E. coli** and have served only for the initial obtention, through PCR or from oligonucleotides, of the several fragments of DNA that constitute the examples. The scheme is self-explanatory, the arrows refer to the steps of cleavage by restriction enzyme, isolation of DNA fragment and its insertion in another plasmid.

Figure 3. Sequence of the DNA fragment encoding a GPI-anchored insulin (Seq. ID No. 1). The complete nucleotide sequence of the construct present on plasmid pBY40, derived from **GAS1** and insulin is shown, along with the amino acid se-

quence of the expected product. In the 5' end, the **GAS1**- derived sequences start at nucleotide 16 and end at nucleotide 380, and in the 3' end, start at nucleotide 560 and end at nucleotide 1058. The insulin-coding sequence starts at nucleotide 393, with a phenylalanine codon, of the B chain, and ends at nucleotide 551, with an asparagine codon of the A chain. Between the B and A chains of insulin, two basic residues were included. The anchor attachment site is at the asparagine residue coded at nucletotide position 649.

Figure 4. Sequence of the DNA fragment encoding a GPI-anchored 18kDa protein (Seq. ID No. 2). The complete nucleotide sequence of the construct present on plasmid pBY48, derived from **GAS1** and the gene encoding 18kDa, is shown. The **GAS1** 5' sequence is identical to that present on plasmid pBY40, shown in Figure 3. From nucleotide 380 to nucleotide 830, the sequence is derived from the gene encoding the 18kDa protein. The remaining sequence is derived from the 3' end of the **GAS1** gene and is identical to that shown in Figure 3.

Figure 5 - Expression of GPI-anchored Human Insulin in Yeast. Colonies of yeast bearing the plasmid pBY40, which codes for GPI-anchored insulin, were grown in a solid medium and transferred, in duplicate, to nitrocellulose filters. As a control, colonies of yeast bearing only the vector YEP352 (here called empty RH273-1A) and colonies bearing a plasmid coding for nomally anchored protein Gas1 were also included. The filters were treated with sodium hydroxide in order to permeabilize the wall of the yeast. In A), the filter was treated with **T. brucei** phospholipase C. In B), the filter was not treated. This was followed by incubation with anti-CRD antibody and visualization of the reaction with protein A marked with iodine-125 after autoradiography.

Figure 6 - Expression of the 18kDa Protein in Anchored Form. Extracts of yeast cells bearing the plasmid pBY48 were prepared by stirring with glass beads. The total extract was subjected to electrophoresis in polyacrylamide gel, in duplicate, and then to immuno-blot. The filter shown in A) was incubated with monoclonal antibody against the 18kDa protein. In B), the filter was treated with **T. brucei** phospholipase C and incubated with anti-CRD antibody. In fil-

ter A), line 1, purified 18kDa protein was applied; in line 2, cellular extract treated with endoglycosidase H; in line 3, extracts of cells lacking the plasmid, treated with endoglycosidase H; in lines 4 and 5, extracts as in 2 and 3 but without the treatment with endoglycosidase H. In filter B), line 1 contains **T. brucei** mfVSG; line 2, extract of empty strain; line 3, extract of strain bearing plasmid with the gene GAS1; line 4, extract of empty strain; line 5, extract of strain expressing anchored 18kDa protein (arrow).

Figure 7 - Solubilization of Anchored Protein by Treatment with phospholipase C. The solubilization can be monitored by phase separation in Triton X-114, followed by the detection of the product with anti-CRD. A sample of extract of wild type **S. cerevisiae** was treated with **B. thuringiensis** PLC (lanes 1) or **T. brucei** PLC (lanes 2) at 30°C for 30 minutes. At the end of the reaction, the concentration of TX-114 was increased to 2% and phase separation was induced by raising the temperature. Duplicates of each phase were subjected to SDS-PAGE, one of them being treated with anti-Gas1 serum (Panel A) and the other with anti-CRD (Panel B). The detection of the bound antibody was made with protein A marked with iodine-125, followed by autoradiography. The detergent and aqueous phases are indicated by D and A, respectively.

The following examples are intended merely to illustrate specific modes of carrying out the present invention and should not be construed as limiting its scope.

Preliminary Comments on the Materials, Strains and Plasmids Reagents for DNA modification and restriction were acquired from Gibco-BRL or BioLabs. Other reagents, salts and materials were acquired from Sigma, Gibco-BRL, Amersham or other similar ones. Media for cultivation of microorganisms were acquired from Difco Laboratories. All the other chemical compounds used are of analytical grade. **E. coli strains** used for the propagation and amplification of the plasmids described herein were DH5 (Hanahan, 1983, J. Mol. Biol. 166:557). The **S. cerevisiae strain** can be RH273-1A, as described in Nuoffer et al; 1991, Mol. Cell. Biol. 11:27. Techniques for handling **S. cerevisiae**, as well as culture media for growing it, and preparation of cellular extracts are de-

scribed in **Guide to Yeast Genetics and Molecular Biology**, Eds. Guthrie & Fink, Meth. Enzymol. **194**. The plasmids described below were obtained in **E. coli** and, after checking their structures by known techniques, they were transferred to **S. cerevisiae**.

Example 1 - Construction of the Vector for Expression of Anchored Proteins. Any plasmid capable of being maintained stable in **S. cerevisiae** cells and having a marker for selection in these cells can be used as a basic vector, as for instance, those containing the gene **LEU2** or **URA3** for complementation of auxotrophic markers, and containing sequences capable of mediating their replication, as derived from "ars" (autonomous replicating sequences) or from the endogenous plasmid 2 micron. They can still contain centromeric sequences that allow the plasmid to be maintained in a low number of copies, in a relatively stable manner. Here one prefers to use the plasmid YEP352 (Hill et al., 1986, Yeast 2:163), which contains the gene **URA3**, and a replication sequence derived from 2 micron so as to mantain it in a high number of copies and, thus, to increase the amount of the recombinant product. In this plasmid a fragment of DNA was inserted which was obtained by a PCR reaction from the cloned GAS1 gene (Nuoffer et al., 1991, Mol. Cell. Biol. 11:27). This fragment will contain the promoter as well as the amino-terminal sequence of **GAS1**, which will allow the transfer of the recombinant polypeptide to the endoplasmic reticulum. This fragment can be obtained through PCR by using, as primers, oligonucleotides complementary, for example, to the region located around the position -250 with respect to the first codon, and to the region located around the position +100 with respect to the GAS1 first codon. In this way, one ensures the presence, in this segment, of the **GAS1** promoter and of the sequence which signals the transport to the endoplasmic reticulum. In this example two primers with the sequences 5'-TTTCCCGGGT-ATTCCTCATACAGC-3' and 5'-ACGGGATCCGTTGGAGTAGAAAAACT-3' were used to obtain a fragment containing the promotor and amino-terminal sequence of the **GAS1** gene, by PCR-amplification of the chromosomal DNA. The 365 nucleotide fragment obtained contain sequences from

248 nucleotides upstream of the start site of translation of the **GAS1** gene up to nucleotide +117 from the starting ATG. The PCR product was blunt-ended cloned into M13, and one clone was chosen for sequencing to determine the correctness of the sequence and for further use. Preferably, this fragment can be inserted in the vector YEp352 so as to contain, downstream of the **GAS1** sequence, a polylinker in order to allow the easy cloning of genes coding for proteins the expression of which is desired. It is convenient to sequence the obtained clone obtained in order to be sure that the sequence is correct. In this plasmid one can still insert a DNA fragment derived from the GAS1 gene by amplification through PCR, coding for the Gas1 protein C terminal sequence and further containing the signal of termination of transcription of the same gene. This fragment can be obtained, for instance, by using a pair of primers complementary to the region close to nucleotide -180 with respect to the **GAS1**-terminating codon, and to the region close to nucleotide +300 with respect to the **GAS1**-terminating codon. In this way, the presence, in this fragment, of the sequence which determines the anchor addition, as well as of sequences of transcription termination and of poly-A tail addition in the message, is ensured. In this example, the carboxyl sequence of the **GAS1** gene, containing the anchor addition site, was obtained by PCR amplification with the oligonucleotides 5'-ACGGTCGACTCTTCTTCCAAGTCTAA-3' and 5'-CCCCAAGCTTGCTGATATTATGGAGAA-3'. The product was blunt-ended cloned initially into BlueScript, for sequencing. The amplified 495 nucleotide fragment encodes the last 61 amino acids of **GAS1**, plus 313 nucleotides of the 3' non-coding region of **GAS1**. The insertion of a fragment coding for any protein "X" whatever between these two Gas1 segments will allow the synthesis of a fusion protein, which after post-translational processings will be anchored to the membrane through a GPI anchor (Figure 1).

Example 2 - Construction of the Vector for Expression of Anchored Insulin. In the plasmid constructed as described above or alternatively as described in Figure 2A, one can insert a chemically or enzymatically synthesized DNA fragment coding for human insulin. For instance, one can ob-

tain a DNA fragment coding for the B and A chains, in this order, and insert, between the amino acids which separate them in the active form, two basic residues which will allow its processing inside the endoplasmic reticulum, or later, after purification of the product. The insulin-encoding sequence, which can be easily derived according to the codon usage frequency for yeast, starting from the known amino acid sequence, can be obtained in several ways, as for instance, through the synthesis of two oligonucleotides so that the nucleotides in the 3' region of one will be complementary to the 3' sequence of the other, in an approximate extent, for instance, of about 20 residues. After annealing, the treatment is effected with a DNA polymerase in order to obtain a double-stranded DNA, which can then be ligated to the previously described vector, in a site suitable for obtaining a fusion with the amino- and carboxy-terminal moieties of Gas1.

In this particular example, the insulin-coding sequence was obtained by the use of two oligonucleotides with the following sequences:
5'-ACGGGATCCAAGAGATTTGTTAACCAACACTTGTGTGGTTCTCACTTGGTTGAAGCCC
TGTACTTGGTTTGTGGTGAAAGAGGTTTCTTCTACACTCCGAAG-3' and
5'-CCGGTCGACTCTGTTACAGTAGTTTTCCAGCTGGTACAAAGAACAAATACTAGTACA
ACATTGTTCAACAATACCTCTCTTAGTCTTCGGAGTGTAGAAGAAAC-3'. After annealing, and polymerization with Klenow, the double-stranded fragment was blunt-end cloned into M13, for sequencing. One particular clone |M13(T2.1)m| was kept for further use. The insulin sequence was then fused to the GAS1-derived sequences, as shown in Figure 2A. Briefly, the sequence derived from the **GAS1** gene encompassing the anchor addition signal, obtained by PCR as described previously, was isolated as a **SalI-HindIII** fragment and inserted in plasmid YEp352, digested with the identical enzymes. This originated plasmid pBY19. A **BamHI-SalI** fragment coding for insulin present in M13(T2.1), obtained as described before, was inserted in pBY19 plasmid, after digestion of the latter with the same enzymes, giving rise to plasmid pBY36. A **BamHI** fragment, derived from plasmid pN#3, containing the promotor and signal sequence of **GAS1**, which was obtained by PCR amplification, was then inserted in the **BamHI** site of plasmid pBY36. The resulting plasmid, pBY40, contains

the necessary information for the synthesis of a GPI-anchored insulin, in yeast. The DNA sequence of this fusion and the amino acid sequence of the resulting protein are given in Figure 3 (Seq. ID N° 1).

Example 3 - Construction of the Vector for Expression of Anchored 18kD Protein. A DNA fragment coding for the **M.** leprae 18 kDa antigen can be obtained by amplification through PCR having as template the cloned gene (Booth et al., 1988, J. Immunol. **140** : 597). The optically synthesized fragment will contain restriction sites for ligation to the vector constructed as described in Example 1 (Figure 2B), so as to allow the correct phase fusion of this sequence to the sequences derived from GAS1.

In this particular example, the sequence encoding the 18kDa protein of **M. leprae** was obtained by PCR amplification using as primers the oligonucleotides 5'-TGCTCTAGACGTACTGACCCGTTCCG-3' and 5'-TGCTCTAGAGGCATCTATGATTTCGT-3', and as template the cloned gene for the 18kDa protein, present in plasmid pUL118. This amplified fragment, containing the entire coding sequence for this protein, was cloned in M13 for sequencing and further use. A **Bam**HI fragment from plasmid pN#3, encompassing the promoter and signal sequence of **GAS1**, was inserted in the **Bam**HI site of plasmid YEp352, originating plasmid pBY32. A **XbaI** fragment containing the coding sequences for the 18kDa protein, present on M13(18K)#5, obtained as described before, was transferred to the **XbaI** site of plasmid pBY32, giving rise to plasmid pBY43. A **SalI-HindIII** fragment containing the **GAS1** C-terminal fragment, present on plasmid pC(B1), was ligated to the **SalI-HindIII** sites of plasmid pBY43, forming plasmid pBY48, which is capable of encoding a GPI-anchored 18kD protein in yeast. The DNA sequence of this construct and the corresponding amino acid sequence of the fusion are given in Figure 4 (Seq. ID N° 2).

Example 4 - Obtention of Yeast Transformed with the above Plasmids. The plasmids obtained in Examples 2 and 3 can be introduced in strains of yeast by conventional methods, such as transformation by lithium acetate or transformation of spheroplasts (**Guide to Yeast Genetics and Molecular Biology,**

Guthrie & Fink, Eds., 1991, Academic Press, Inc.). The strain to be used can be preferably RH273-1A, since it lacks the Gas1 protein. Transformants should be selected in a minimum medium through selection by complementation of auxotrophic markers, in this case growth in the absence of uracil.

In this example, the strain constructed to express GPI-anchored insulin is called RH273-1A/pBY40, and the strain constructed to express GPI-anchored 18kDa protein is called RH273-1A/pBY48.

Example 5 - Detection of the Expression of GPI-Anchored Insulin. Strain RH273-1A/pBY40, constructed as described in Example 4, can be grown in isolated colonies, initially in a selective medium, and thereafter replicated to a rich medium. The cells can be transferred to nitrocellulose membranes, lysed, treated with phospholipase C obtained from **Trypanosoma brucei**, and with anti-CRD antiibody. The presence of CRD epitope can be visualized through the incubation of this filter with protein A coupled to iodine-125, as described in Cardoso de Almeida & Turner, 1983, op.cit. (figure 5).

Example 6 - Detection of the Expression of GPI-Anchored 18kDa protein. Similarly as described in Example 5, the expression of anchored 18 kDa protein in RH273-1A/pBY48 cells can be detected by using the anti-CRD antibody. Alternatively, the anchored 18 kDa protein can be easily visualized in immuno-blots from SDA-PAGE of cellular extracts (figure 6). Both monoclonal L5 and anti-CRD antibodies recognize the recombinant product, which is highly glycosylated, as shown by the high molecular weight of the product. The glycosylation occurs in residues derived from Gas1, and can be eliminated by the construction of a fusion wherein the residues susceptible of glycosylation are removed.

Example 7 - Release with GPI-specific phospholipase. The GPI-anchored proteins can be solubilized by treating the cells or cellular extracts with phospholipase specific for GPI structures, as shown in figure 7.

Example 8. Utilization of GPI as a vaccine or immune system modulator. The GPI moieties produced by engineered organism can be purified and analysed according to standard procedures such as solvent selective extraction and

fractionation by thin layer chromatography or reverse phase thin layer chromatography or by cromatography on Octyl-Sepharose. These glycolipids, or part of them, chemically or enzymatically treated, are then prepared with appropriate adjuvants and administered as vaccines or immune system modulators. Their effects can be followed by classical methods of humoral and cellular immunology such as titrating levels of antibodies with certain specificities or by analysing the profile of lymphocytes which can be stimulated by the injected antigen.

The sequences shown in figures 3 and 4, Seq. ID No. 1 and Seq. ID No. 2, without however the amino acid indication, are also submitted separately both in printed and in computer readable form.

CLAIMS

1. A process for producing a recombinant protein, or a precursor thereof, in cells of genetically modified eukaryotic microorganisms, characterized by comprising the steps of:

a) biosynthesizing said protein or a precursor thereof by the cell of said microorganism and linking the endogenous glycosylphosphatidylinositol (GPI) to the C-terminal amino acid of the obtained protein or precursor, with the consequent anchorage of said protein or precursor to membranes of said microorganism by means of GPI; and

b) selectively releasing said protein or precursor, obtained in step (a), by methods making use of intrinsic properties provided by the presence of GPI.

2. A process according to claim 1, characterized in that said eukaryotic microorganism is a genetically modified **Saccharomyces cerevisiae.**

3. A process according to claim 2, characterized in that said eukaryotic microorganism is a **Saccharomyces cerevisiae**, genetically modified so as not to produce GPI-anchored endogenous proteins.

4. A process according to claim 2, characterized in that said eukaryotic microorganism is **Saccharomyces cerevisiae**, genetically modified so as not to produce the endogenous protein Gasl.

5. A process according to any one of claims 1 to 4, characterized in that said recombinant protein is human insulin.

6. Process according to any one of claims 1 to 4, characterized in that said recombinant protein is **Mycobacterium leprae** antigen 18kDa.

7. A process according to any one of claims 1 to 6, characterized in that the selective release of the protein or its precursor obtained in step a) is carried out by enzymatic treatment.

8. A process according to claim 7, characterized in that the treatment is carried out with PI- or GPI-specific

phospholipase.

9. A process according to claim 8, characterized in that the treatment is carried out with PI- or GPI-specific phospholipase of C or D specificity.

10. A process according to claim 7, characterized in that the treatment is carried out with proteases.

11. A process according to any one of claim 1 to 6, characterized in that the selective release of the protein or its precursor obtained in step a) is carried out by chemical treatment.

12. A process according to claim 11, characterized in that the treatment is carried out by nitrous deamination.

13. A process according to claim 11, characterized in that the treatment is carried out with a base.

14. A process according to any one of claims 1 to 6, characterized in that the selective release of the protein or its precursor obtained in step a) is carried out by treatment with neutral detergents.

15. A process according to claim 14, characterized in that the treatment is carried out with neutral detergents of the type n-octyl glucopyranoside.

16. A process according to any of claims 1 to 6, characterized in that the selective release of the protein or its precursor obtained in step a) is carried out by co-expression of a PI- or GPI-specific phospholipase in the cell of said microorganism.

17. A process according to claim 9, characterized in that one works with a PI- or GPI-specific phospholipase C, thereby obtaining hydrolysed protein or its precursor, which has in its structure cyclic 1,2-phosphate inositol (epitope CRD), which is susceptible of being selectively purified by immunoafinity with anti-CRD antibodies.

18. A process for the production of glycosylphosphatidylinositol (GPI) and of a recombinant protein in cells of genetically modified eukaryotic microorganisms, characterized by comprising the steps of:

a) biosynthesising said protein or a precursor thereof by the cell of said microorganism and linking the endogenous glycosylphosphatidylinositol (GPI) to the C-

terminal amino acid of the obtained protein or precursor, with the consequent anchorage of said protein or precursor to membranes of said microorganism by means of GPI;

b) selectively releasing said protein or precursor, obtained in step (a), by methods making use of intrinsic properties provided by the presence of GPI, and

c) selectively releasing GPI.

19. A process according to claim 18, characterized in that said eukaryotic microorganism is a genetically modified **Saccharomyces cerevisiae.**

20. A process according to claim 19, characterized in that said eukaryotic microorganism is a **Saccharomyces cerevisiae,** genetically modified so as not to produce endogenous GPI-anchored proteins.

21. A process according to claim 19, characterized in that said eukaryotic microorganism is a **Sacchromyces rerevisiae,** genetically modified so as not to produce the endogenous protein **Gas1.**

22. A process according to any one of claim 18 to 21, characterized in that said recombinant protein is human insulin.

23. A process according to any one of claims 18 to 21, characterized in that said recombinant protein is **Mycobacterium leprae** antigen 18kDa.

24. A process according to any one of claims 18 to 23, characterized in that the selective release of the protein or its precursor or GPI or of its precursor obtained in step a) is carried out by enzymatic treatment.

25. A process according to claim 24, characterized in that the treatment is carried out with PI- or GPI-specific phospholipase.

26. A process according to claim 25, characterized in that the treatment is carried out with PI- or GPI-specific phospholipase of C or D specificity.

27. A process according to claim 24, characterized in that the treatment is carried out with proteases.

28. A process according to any one of claims 18 to 23, characterized in that the selective release of the protein or its precursor or of GPI or its precursor obtained in step

a) is carried out by chemical treatment.

29. A process according to claim 28, characterized in that the treatment comprises extractions with organic solvents.

30. A process according to claim 28, characterized in that the treatment is carried out with a base.

31. A process according to any one of claims 18 to 23, characterized in that the selective release of the protein or its precursor obtained in step a) is carried out by treatment with neutral detergents.

32. A process according to claim 31, characterized in that the treatment is carried out with neutral detergents of the type n-octyl glucopyranoside.

33. A process according to any one of claims 18 to 23, characterized in that the selective release of the protein or its precursor obtained in step a) is carried out by co-expression of a PI- or GPI-specific phospholipase in the cell of said microorganism.

34. A process according to claim 26, characterized in that one works with PI- or GPI-specific phospholipase C, thereby obtaining hydrolysed protein or its precursor, which has in its structure cyclic 1,2-phosphate inositol (epitope CRD), which is susceptible of being selectively purified by immunoafinity with anti-CRD antibodies.

35. A process for the production of glycosylphosphatidylinositol (GPI) in cells of a genetically modified **Saccharomyces cerevisiae** characterized by comprising the steps of:

a) biosynthesis of a protein or a precursor thereof and biosynthesis of GPI or a precursor thereof by said cells ; and

b) selective release of GPI.

36. A process according to claim 36, characterized in that said **Saccharomyces cerevisiae** is one genetically modified so as not to produce the endogenous protein Gas1.

37. A process for the obtention of cells of **Saccharomyces cerevisiae** yeasts capable of expressing a recombinant protein, characterized by comprising the steps of:

a) providing the gene coding for said recombinant

protein or its precursor so as to be anchored via GPI;

b) inserting the gene, obtained in step a), in an suitable vector;

c) introducing the vector containing the gene coding for said recombinant protein or its precursor in a **Sacchromyces cerevisiae** cell; and

d) propagating the cell obtained in step c).

38. A process according to claim 37, characterized in that said recombinant protein is human insulin.

39. A process according to claim 37, characterized in that said recombinant protein is **Mycobacterium leprae** 18kDa antigen.

40. A process according to claim 37, characterized in that the **Saccharomyces cerevisiae** cell used lacks the capability of producing the endogenous protein **Gasl**.

41. A yeast cell, characterized by being genetically modified so as to express a GPI-anchored recombinant protein.

42. A cell according to claim 41, characterized by being **Saccharomyces cerevisiae.**

43. A cell according to claim 42, characterized by not expressing the endogenous protein **Gasl**.

44. A cell according to any one of claims 40 to 43, characterized in that said recombinant protein is human insulin.

45. A cell according to any one of claims 40 to 43, characterized in that said recombinant protein is **Micobacterium leprae** 18kDa antigen.

46. A nucleotide sequence, characterized by coding for a recombinant protein or its precursor, susceptible of being anchored by GPI.

47. A nucleotide sequence, characterized by coding for:

i) the C-terminal moiety of the protein **Gasl** of **Saccharomyces cerevisiae**;

(ii) a protein or a precursor thereof, and

(iii) the N-terminal moiety of the protein **Gasl** of **Saccharomyces cerevisiae.**

48. A nucleotide sequence according to claim 47, characterized in that said protein is human insulin.

49. A nucleotide sequence according to claim 47, characterized in that said protein is **Mycobacterium leprae** 18kDa antigen.

50. A culture medium, characterized by containing cells according to any one of claim s 41 to 43.

51. A medicine or vaccine, characterized by containing a recombinant protein obtained by a process according to any one of claims 1 to 34 or a recombinant protein of which the precursor was obtained by a process according to any one of claims 1 to 34.

52. A medicine according to claim 51, characterized by containing recombinant human insulin.

53. A vaccine according to claim 51, characterized by containing **Mycobacterium leprae** 18kDa antigen.

54. A medicine or vaccine characterized by containing a GPI obtained by a process according to any one of claims 35 to 40 or a GPI of which the precursor was obtained by a process according to any one of claims 35 to 40.

55. A product obtainable by a process according to any one of claims 1 to 40.

56. A process according to claim 35, characterized in that step a) further comprises linking the endogenous glycosylphosphatydylinositol (GPI) to the C-terminal amino acid of the obtained protein or its precursor, with the consequent anchoring of said protein or precursor to membranes of said microorganism by means of GPI.

57. A cell according to any one of claims 41 - 45, characterized by the capacity of increasing the levels of GPI biosynthesis or of producing GPI with determined properties.

58. Method for detecting the presence of a GPI anchor in insulin or in 18kDa protein, by means of immuno-blots with antibody anti-CRD after treatment of the filters with PLC.

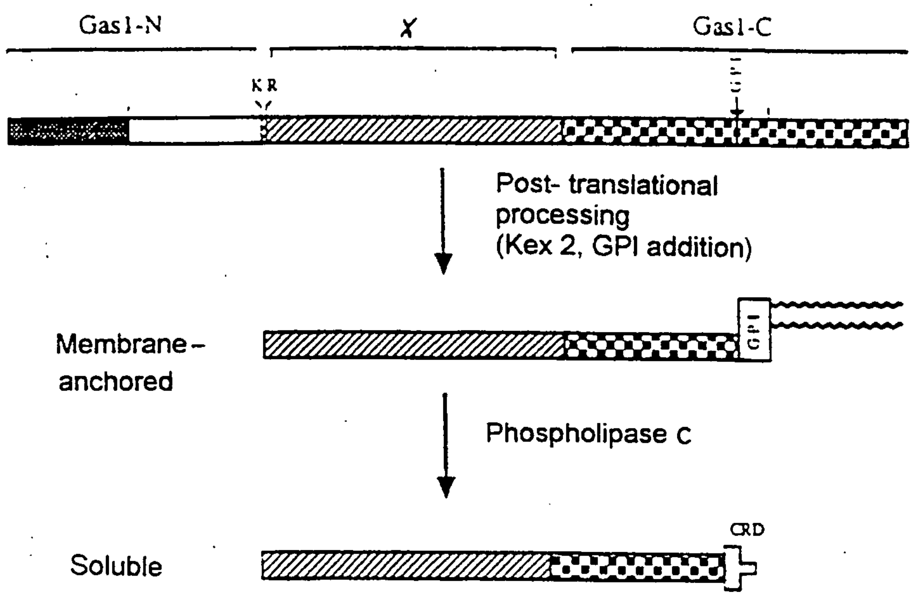


FIGURE 1. Scheme of the precursor and of the GPI - anchored final product in S. cerevisiae, and of its release with phospholipase

FIGURE 2A.

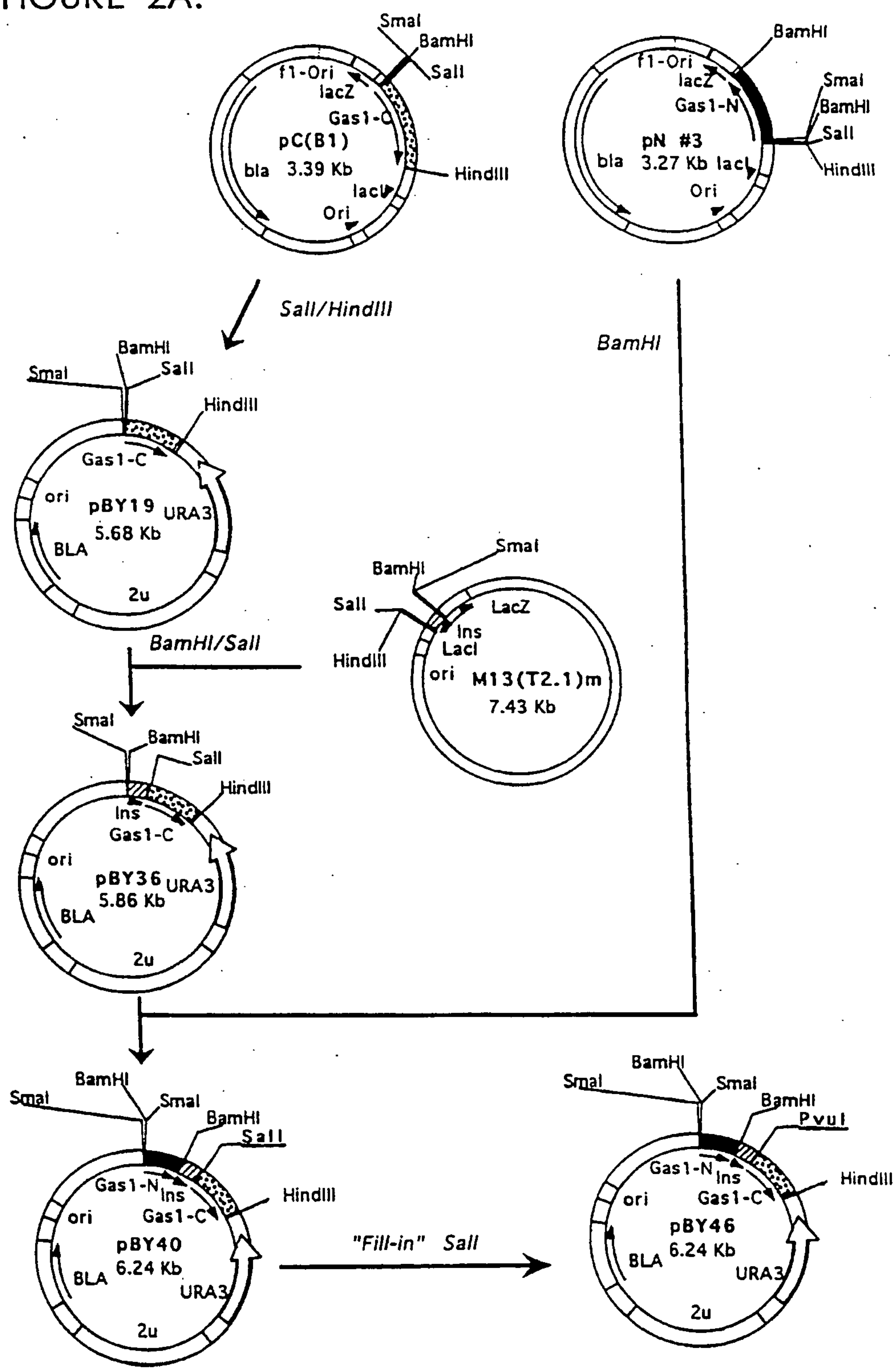

SmaI
BamHI
SalI
f1-Ori
lacZ
Gas1-C
pC(B1)
bla 3.39 Kb
HindIII
lacI
Ori
BamHI
f1-Ori
lacZ
Gas1-N
SmaI
BamHI
SalI
pN #3
bla 3.27 Kb lacI
HindIII
Ori
SalI/HindIII
BamHI
BamHI
SmaI
SalI
HindIII
Gas1-C
ori
pBY19
5.68 Kb
URA3
BLA
2u
SmaI
BamHI
SalI
LacZ
Ins
LacI
HindIII
ori
M13(T2.1)m
7.43 Kb
BamHI/SalI
SmaI
BamHI
SalI
HindIII
Ins
Gas1-C
ori
pBY36
5.86 Kb
URA3
BLA
2u
BamHI
SmaI
SmaI
BamHI
SalI
Gas1-N
Ins
Gas1-C
HindIII
ori
pBY40
6.24 Kb
BLA
URA3
2u
"Fill-in" SalI
BamHI
SmaI
SmaI
BamHI
PvuI
Gas1-N
Ins
Gas1-C
HindIII
ori
pBY46
6.24 Kb
BLA
URA3
2u

FIGURE 2B.

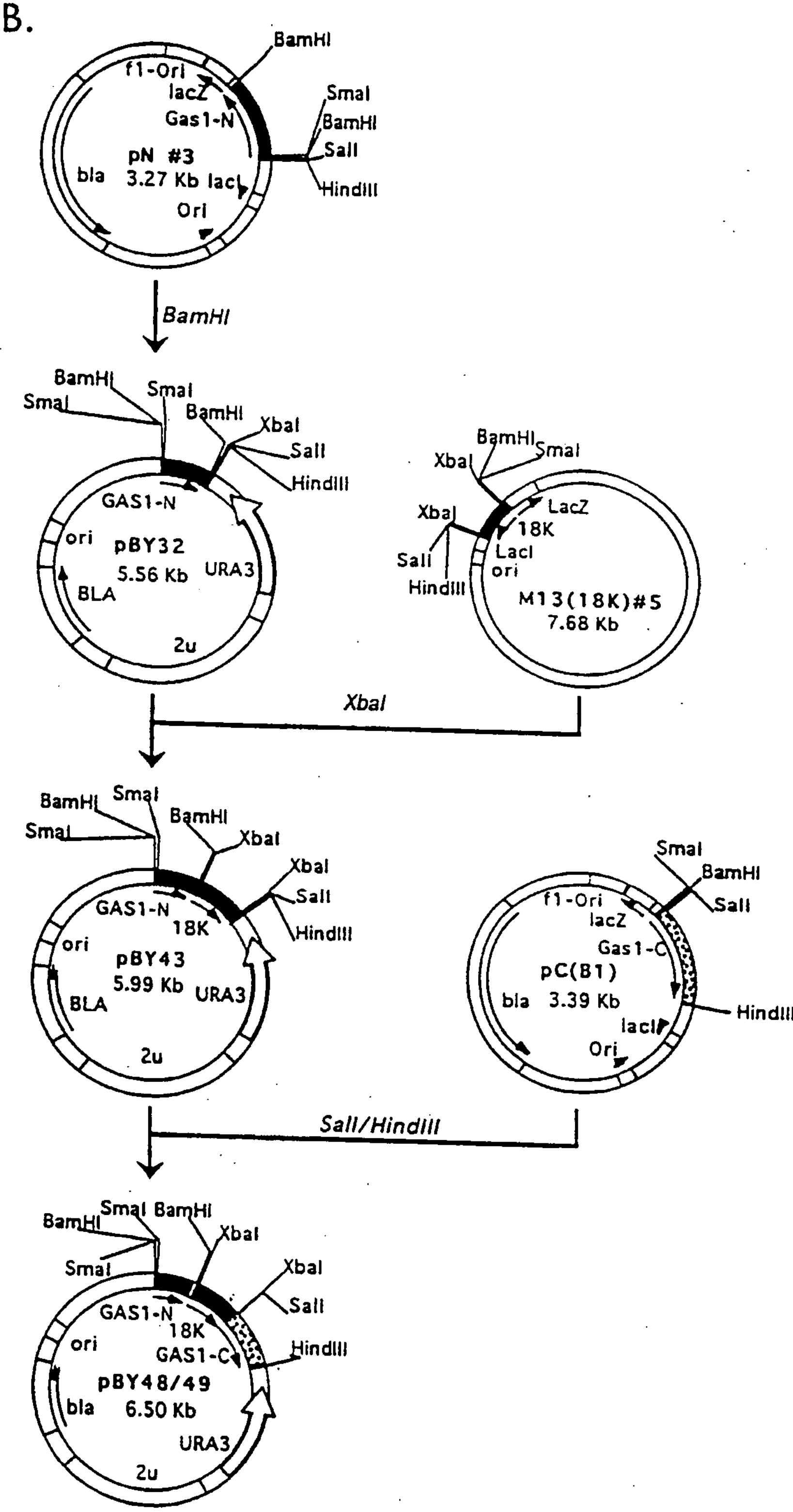

BamHI
SmaI
BamHI
SalI
HindIII
f1-Ori
lacZ
Gas1-N
pN #3
3.27 Kb
bla
lacI
Ori
BamHI
SmaI
BamHI
SmaI
BamHI
XbaI
SalI
HindIII
GAS1-N
ori
pBY32
5.56 Kb
URA3
BLA
2u
BamHI
SmaI
XbaI
XbaI
SalI
HindIII
LacZ
18K
LacI
ori
M13(18K)#5
7.68 Kb
XbaI
BamHI
SmaI
SmaI
BamHI
XbaI
XbaI
SalI
HindIII
GAS1-N
18K
ori
pBY43
5.99 Kb
URA3
BLA
2u
SmaI
BamHI
SalI
f1-Ori
lacZ
Gas1-C
pC(B1)
bla
3.39 Kb
HindIII
lacI
Ori
SalI/HindIII
BamHI
SmaI
BamHI
SmaI
XbaI
XbaI
SalI
HindIII
GAS1-N
18K
GAS1-C
ori
pBY48/49
bla
6.50 Kb
URA3
2u

FIGURE 3

ANCHORED INSULIN (pBY40) SEQUENCE

```
GGATCCCCTTTCCCGGGTATTCCTCATACAGCCTGCGCGGTTTATTAGTAAAATACCCGA
         10        20        30        40        50        60

TAATCCTCGAGGTTTGAAAAACTTTTCCCTCTACTACTGTTGACACGGATTTTTTTATTT
         70        80        90       100       110       120

AAGAGGAAAAGTCGTGGTTGTTTTCCTCGAACAAATTAGATATCCATAAATAGTTGTGTC
        130       140       150       160       170       180

GTTTTATTAAGCTATTTCAAAATCAGTTTTTATTTTTTAAAGTCTGATAAAACAAAAACA
        190       200       210       220       230       240

                         M  L  F  K  S  L  S  K  L  A  T  A
ACAAACACAGCTAAATCTCAACAATGTTGTTTAAATCCCTTTCAAAGTTAGCAACCGCTG
        250       260       270       280       290       300

A  A  F  F  A  G  V  A  T  A  D  D  V  P  A  I  E  V  V  G
CTGCTTTTTTTGCTGGCGTCGCAACTGCGGACGATGTTCCAGCGATTGAAGTTGTTGGTA
        310       320       330       340       350       360

N  K  F  F  Y  S  N  G  S  K  R  F  V  N  Q  H  L  C  G  S
ATAAGTTTTTCTACTCCAACGGATCCAAGAGATTTGTTAACCAACACTTGTGTGGTTCTC
        370       380       390       400       410       420

H  L  V  E  A  L  Y  L  V  C  G  E  R  G  F  F  Y  T  P  K
ACTTGGTTGAAGCCCTGTACTTGGTTTGTGGTGAAAGAGGTTTCTTCTACACTCCGAAGA
        430       440       450       460       470       480

T  K  R  G  I  V  E  Q  C  C  T  S  I  C  S  L  Y  Q  L  E
CTAAGAGAGGTATTGTTGAACAATGTTGTACTAGTATTTGTTCTTTGTACCAGCTGGAAA
        490       500       510       520       530       540

N  Y  C  N  R  V  D  S  S  S  K  S  N  S  G  S  S  G  S  S
ACTACTGTAACAGAGTCGACTCTTCTTCCAAGTCTAACTCCGGCTCTTCTGGTTCTTCCA
        550       560       570       580       590       600

S  S  S  S  S  S  S  A  S  S  S  S  S  S  K  K  N  A  A  T
GTTCTTCTTCTTCTTCTTCAGCTTCATCTTCATCTTCTAGCAAGAAGAATGCTGCCACCA
        610       620       630       640       650       660
```

FIGURE 3 **(cont)**

```
N  V  K  A  N  L  A  Q  V  V  F  T  S  I  I  S  L  S  I  A
ACGTTAAAGCTAACTTAGCACAAGTGGTCTTTACCTCCATCATTTCCTTATCCATTGCCG
         670       680       690       700       710       720

A  G  V  G  F  A  L  V  *
CTGGTGTCGGTTTTGCTTTGGTTTAAAAAGCTagctTCGACACATACATAATAACTCGAT
         730       740       750       760       770       780

AAGGTATGGTATCTTATTTCATTGTGGGTAGTTTTTTACGAAAAAAATGAAAAGTTGTAA
         790       800       810       820       830       840

GTATAGTATATATTTTTTTTCTATGTAAGTTTTATAAGATTCTATTCGCTATTACCACCG
         850       860       870       880       890       900

GTAAATTAAAAAGAACACTATTGTTACATTATATGTTTTTAAATCATCAAAAAAGACAAT
         910       920       930       940       950       960

ATTCATTTAATATTCCTTATAGAACTACTTAACATTGTTCTTCTTTCTATTAAACGTCTT
         970       980       990      1000      1010      1020

TATGCAAACCATTTATGTACTTTCTCCATAATATCAGCAAGCTT
        1030      1040      1050      1060
```

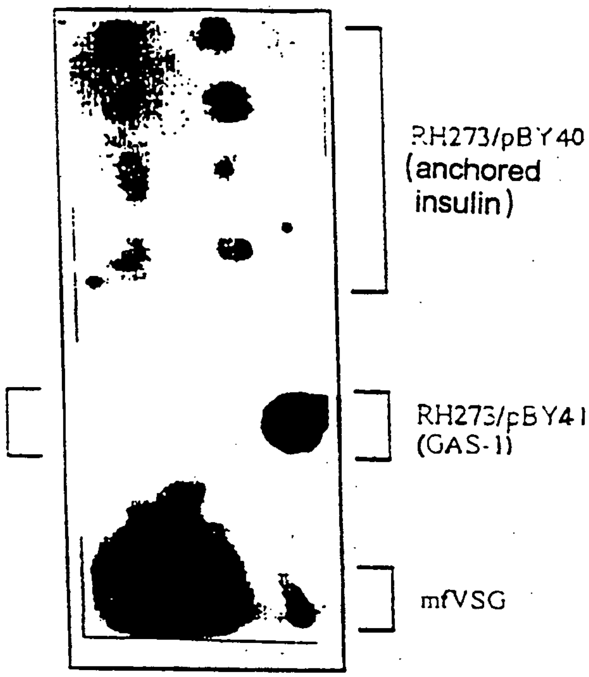
FIGURE 4

ANCHORED 18kD (pBY48/49) SEQUENCE

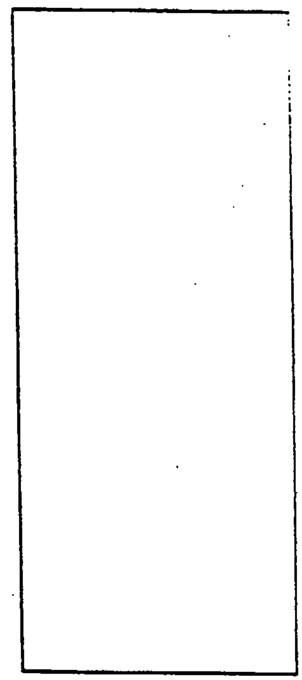
```
GGATCCCCTTTCCCGGGTATTCCTCATACAGCCTGCGCGGTTTATTAGTAAAATACCCGA
         10        20        30        40        50        60

TAATCCTCGAGGTTTGAAAAACTTTTCCCTCTACTACTGTTGACACGGATTTTTTTATTT
         70        80        90       100       110       120

AAGAGGAAAAGTCGTGGTTGTTTTCCTCGAACAAATTAGATATCCATAAATAGTTGTGTC
        130       140       150       160       170       180

GTTTTATTAAGCTATTTCAAAATCAGTTTTTATTTTTTAAAGTCTGATAAAACAAAAACA
        190       200       210       220       230       240

                        M  L  F  K  S  L  S  K  L  A  T  A
ACAAACACAGCTAAATCTCAACAATGTTGTTTAAATCCCTTTCAAAGTTAGCAACCGCTG
        250       260       270       280       290       300

A  A  F  F  A  G  V  A  T  A  D  D  V  P  A  I  E  V  V  G
CTGCTTTTTTTGCTGGCGTCGCAACTGCGGACGATGTTCCAGCGATTGAAGTTGTTGGTA
        310       320       330       340       350       360

N  K  F  F  Y  S  N  G  S  S  R  R  T  D  P  F  R  E  L  D
ATAAGTTTTTCTACTCCAACGGATCCTCTAGACGTACTGACCCGTTCCGTGAACTGGACC
        370       380       390       400       410       420

R  F  A  E  Q  V  L  G  T  S  A  R  P  A  V  M  P  M  D  A
GCTTCGCCGAGCAAGTGTTAGGTACGTCTGCCCGCCCAGCAGTAATGCCCATGGACGCTT
        430       440       450       460       470       480

W  R  E  G  E  E  F  V  V  E  F  D  L  P  G  I  K  A  D  S
GGCGTGAGGGCGAAGAATTCGTCGTCGAGTTCGACCTTCCTGGCATCAAAGCCGATTCAC
        490       500       510       520       530       540

L  D  I  D  I  E  R  N  V  V  T  V  R  A  E  R  P  G  V  D
TGGACATTGACATCGAACGCAACGTAGTCACCGTGCGGGCCGAGCGCCCAGGCGTCGACC
        550       560       570       580       590       600

P  D  R  E  M  L  A  A  E  R  P  R  G  V  F  N  R  Q  L  V
CCGATCGGGAAATGCTTGCTGCCGAGCGGCCACGCGGTGTGTTCAATCGTCAGCTGGTTC
        610       620       630       640       650       660
```

FIGURE 4 (cont)

```
L  G  E  N  L  D  T  E  R  I  L  A  S  Y  Q  E  G  V  L  K
TCGGCGAAAACCTCGACACCGAACGGATCTTGGCTTCCTACCAAGAAGGTGTCCTGAAGT
        670       680       690       700       710       720

L  S  I  P  V  A  E  R  A  K  P  R  K  I  S  V  D  R  G  N
TGTCGATACCAGTAGCCGAAAGGGCTAAACCGCGCAAGATCTCCGTTGATCGTGGCAACA
        730       740       750       760       770       780

N  G  H  Q  T  I  N  K  T  A  H  E  I  I  D  A  S  R  V  D
ACGGACACCAGACCATAAACAAAACCGCACACGAAATCATAGATGCCTCTAGAGTCGACT
        790       800       810       820       830       840

S  S  S  K  S  N  S  G  S  S  G  S  S  S  S  S  S  S  S  S
CTTCTTCCAAGTCTAACTCCGGCTCTTCTGGTTCTTCCAGTTCTTCTTCTTCTTCTTCAG
        850       860       870       880       890       900

A  S  S  S  S  S  S  K  K  N  A  A  T  N  V  K  A  N  L  A
CTTCATCTTCATCTTCTAGCAAGAAGAATGCTGCCACCAACGTTAAAGCTAACTTAGCAC
        910       920       930       940       950       960

Q  V  V  F  T  S  I  I  S  L  S  I  A  A  G  V  G  F  A  L
AAGTGGTCTTTACCTCCATCATTTCCTTATCCATTGCCGCTGGTGTCGGTTTTGCTTTGG
        970       980       990      1000      1010      1020

V  *
TTTAAAAAGCTagctTCGACACATACATAATAACTCGATAAGGTATGGTATCTTATTTCA
       1030      1040      1050      1060      1070      1080

TTGTGGGGTAGTTTTTACGAAAAAAAATGAAAAGTTGTAAGTATAGTATATATTTTTTTTC
       1090      1100      1110      1120      1130      1140

TATGTAAGTTTTATAAGATTCTATTCGCTATTACCACCGGTAAATTAAAAAGAACACTAT
       1150      1160      1170      1180      1190      1200

TGTTACATTATATGTTTTTAAATCATCAAAAAAGACAATATTCATTTAATATTCCTTATA
       1210      1220      1230      1240      1250      1260

GAACTACTTAACATTGTTCTTCTTTCTATTAAACGTCTTTATGCAAACCATTTATGTACT
       1270      1280      1290      1300      1310      1320

TTCTCCATAATATCAGCAAGCTT
       1330      1340
```



A) +Tb PLC



B) -Tb PLC

FIGURE 5. Expression of GPI - anchored human insulin in yeast

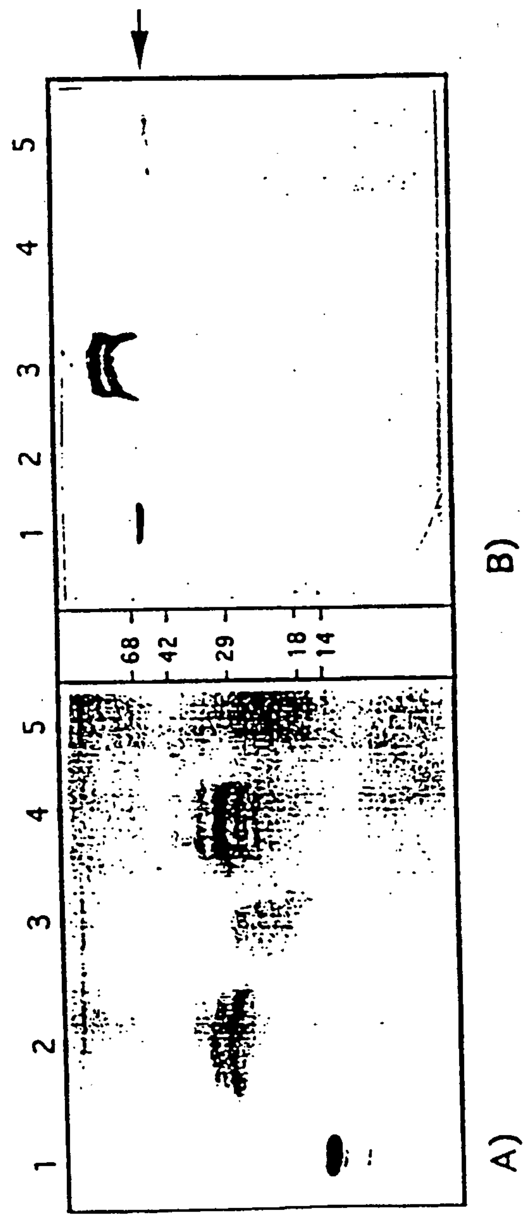


FIGURE 6. Expression of 18kDa protein in anchored form

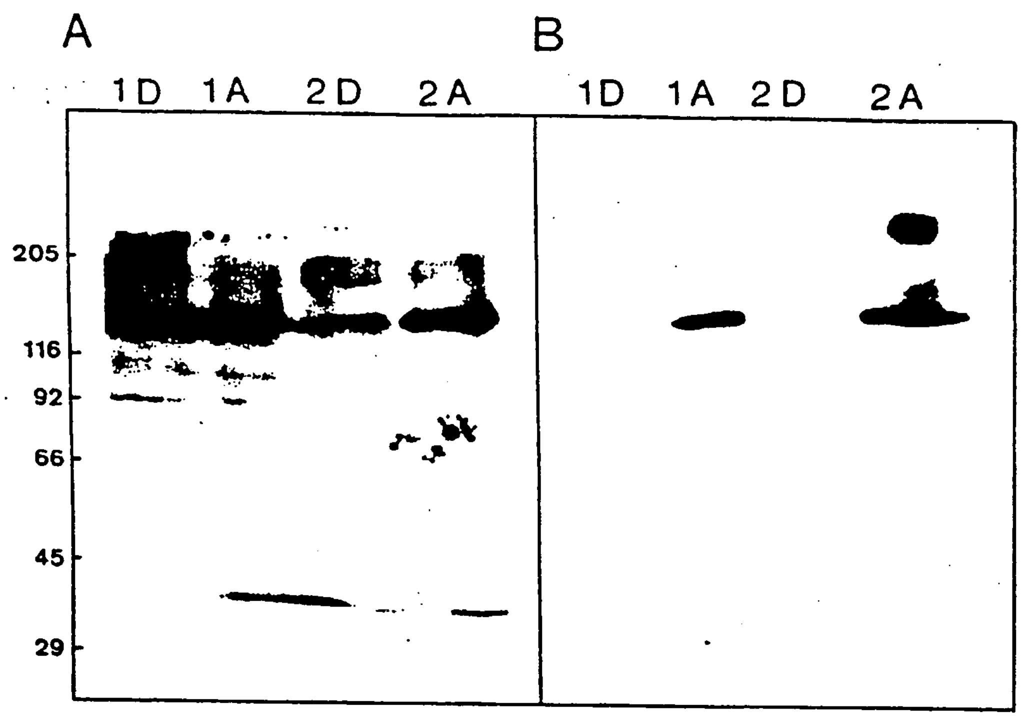


FIGURE 7. Solubilization of anchored protein by treatment with phospholipase C.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/BR 95/00010

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/17 C12N15/31 C12N15/62 C12P7/64 C12N1/19 C12N5/10 A61K38/00 A61K38/28 A61K39/04 C07K14/62 C07K14/35 C07H13/06 G01N33/74 G01N33/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K C12N C12P A61K C07H G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO,A,94 01463 (UNIVERSITY OF BRITISH COLUMBIA) 20 January 1994 see page 10, line 31 - page 11, line 20 see page 22, line 13 - line 32 see page 23, line 6 - line 30 see page 28, line 13 - page 29, line 8 see page 29, line 35 - page 30, line 13 see page 32, line 13 - page 33, line 25 ---	1,2,7-9, 37,41, 42,46, 50,51,55
X	EP,A,0 477 739 (F. HOFFMANN-LA ROCHE AG) 1 April 1992 see page 2, line 32 - line 38; examples 3-6 --- -/--	1,7-9, 16,46, 51,55

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
23 June 1995	04.07.95

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Montero Lopez, B

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO,A,92 01715 (UNIV LELAND STANFORD JUNIOR) 6 February 1992 see page 3, line 21 - line 37 see page 4, line 21 - page 5, line 36 see page 6, line 10 - page 8, line 34 ---	1,7-12, 46,55
X	WO,A,90 12099 (BIOGEN INC) 18 October 1990 see page 1, line 3 - line 21 see page 6, line 34 - page 7, line 12 see page 8, line 22 - page 10, line 19 see page 10, line 34 - page 12, line 35 see page 13, line 34 - page 14, line 29 ---	1,7-12, 46,55
X	WO,A,89 01041 (GENENTECH INC) 9 February 1989 see page 5, line 20 - line 32; example 3 see page 10, line 31 - page 11, line 16 see page 12, line 1 - line 18 see page 13, line 18 - page 15, line 33 see page 24, line 29 - line 34 ---	1,2,37, 41,42, 46,50, 51,55
X	DATABASE WPI Section Ch, Week 9344 Derwent Publications Ltd., London, GB; Class B05, AN 93-351329 BOOTH, ROGER, J. ET AL. 'Liposome composition containing phosphatidyl inositol and optionally phosphatidyl serine. ' & WO,A,93 20801 (HISAMITSU PHARM CO LTD) , 28 October 1993 see abstract ---	54
X	WO,A,89 05825 (IMMUNEX CORP) 29 June 1989 see page 2, line 34 - page 3, line 4 ---	51,53
X	EP,A,0 055 885 (LILLY CO ELI) 14 July 1982 see page 1, line 1 - line 15 ---	52
A	EUROPEAN JOURNAL OF BIOCHEMISTRY, vol. 176, no. 3, October 1988 pages 527-534, SUSANNE E. ZAMZE ET AL. 'Characterization of thr cross-reacting determinant (CRD) of the glycosyl-phosphatidylinositol membrane anchor of Trypanosoma brucei variant surface glycoprotein' cited in the application see abstract see page 532, right column, paragraph 2 - page 533, right column, paragraph 3 --- -/--	58

1

In ational Application No
PCT/BR 95/00010

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	BRAZ. J. MED. BIOL. RES. (1994), 27(3), 623-6 CODEN: BJMRDK;ISSN: 0100-879X, vol. 27, no. 3, March 1994 pages 623-626, AMORIM, A. G. ET AL 'Expression of Mycobacterium leprae 18-kDa antigen in yeast in a GPI -anchored form' see abstract see page 624, paragraph 2 - page 625, paragraph 3; figure 1 -----	1-4,6-9, 17,37, 39-43, 45-50,55

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

In tional Application No
PCT/BR 95/00010

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9401463	20-01-94	AU-B- 4553793	31-01-94
		EP-A- 0651768	10-05-95
EP-A-477739	01-04-92	JP-A- 5076357	30-03-93
		US-A- 5418147	23-05-95
WO-A-9201715	06-02-92	NONE	
WO-A-9012099	18-10-90	US-A- 5223394	29-06-93
		AU-A- 5428990	05-11-90
		CA-A- 2051694	11-10-90
		EP-A- 0466806	22-01-92
		JP-T- 4504360	06-08-92
WO-A-8901041	09-02-89	US-A- 5109113	28-04-92
		AU-B- 629517	08-10-92
		AU-A- 2308788	01-03-89
		EP-A- 0371999	13-06-90
		JP-T- 2504467	20-12-90
		US-A- 5374548	20-12-94
		US-A- 5264357	23-11-93
WO-A-8905825	29-06-89	NONE	
EP-A-0055885	14-07-82	NONE	